4-17-2015

Gradient Effects of Reading Ability on Talker-Specific Perceptual Learning

Shayna R. Marmon
University of Connecticut - Storrs, shayna.marmon@uconn.edu

Gradient Effects of Reading Ability on Talker-Specific Perceptual Learning

Shayna Rebecca Marmon

B.A., University of Connecticut, 2013

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

At the

University of Connecticut

2015

APPROVAL PAGE

Master of Arts Thesis

Gradient Effects of Reading Ability on Talker-Specific Perceptual Learning

Presented by

Shayna Rebecca Marmon

Major Advisor ____

Rachel M. Theodore, Ph.D.

Associate Advisor ____

Emily B. Myers, Ph.D.

Associate Advisor ____

Pradeep Ramanathan, Ph.D.

University of Connecticut

2015

Acknowledgments

I would like to thank Dr. Rachel M. Theodore for her dedication and support towards the completion of this Master's Thesis. Her endless guidance over the past three years has fostered my passion for research and asking questions that I will take with me throughout my professional future.

I would like to acknowledge committee members Dr. Emily B. Myers and Dr. Pradeep Ramanathan. I would also like to thank Ms. Katlyn Salvador for her significant contributions towards stimuli development, data collection, and collaborative support that contributed to the completion of the current work. I also extend thanks to undergraduate research assistants Stephen Graham, SarahAnn Cooper, Shannon Hughes, Jenean Jourdan, William Baron, Carol Baingana, Ashley Bean, John Gerrity, Jr., and Madeline Moulton for assistance with data collection.

This research was supported by the UConn Speech and Hearing Clinic [Clinical Research Grant to R. Theodore]; the College of Liberal Arts and Sciences at the University of Connecticut [start-up funds to R. Theodore]; the UConn Office of Undergraduate Research [Social Sciences, Humanities, and Arts Research Experience Award to K. Salvador; Summer Undergraduate Research Award to K. Salvador]; and the Department of Psychology at the University of Connecticut [Undergraduate Research Grant to K. Salvador].

Table of Contents

Abstract

The acoustic signal of speech is a complex signal that simultaneously cues the linguistic content of a talker's message and the identity of the talker. Traditional models of speech perception proposed that the properties of the acoustic signal that cue linguistic meaning and talker identify were distinct. However, a growing body of evidence challenges this traditional view, indicating that linguistic and talker-specific phonetic information are tightly linked in language comprehension. One of these benefits is increased intelligibility of the spoken message when listening to a familiar talker compared to an unfamiliar talker. Just as experience with a talker's voice influence linguistic processing, it has been shown that this link is bi-directional such that linguistic experience influence voice recognition ability, indicating the recruitment of phonological knowledge to perform talker identification tasks. Consistent with this account, recent evidence indicates that adults with dyslexia demonstrate reduced talker identification abilities when compared to a typically developing population. Other work suggests that there may be a gradient influence of phonological ability on talker recognition. Here we test the hypothesis that reading ability will demonstrate a gradient effect on the ability to not only learn talkers' voices, but also the ability to incorporate talker-specific phonetic detail into lexical representations across the range abilities in the unimpaired population. Monolingual English readers were assigned to either the average (n = 15) or advanced (n = 15) reading group based on a median split of performance on a standardized assessment battery for reading sub-skills and comprehension. All readers participated in a five-session study, where at session 1 each completed a word transcription pre-test, sessions 2 – 4 consisted of talker identification training for six talkers, and session 5 consistent of a post-test identical to the transcription test used in session 1. The results indicated that compared to the average readers, the advanced readers (1) showed higher rate of learning for talker identification during training and (2) showed greater degree of improvement on speech perception at post-test. Correlational analysis showed that the degree of improvement at post-test was significantly correlated with performance on the standardized reading assessment battery. These results extend earlier findings to include a gradient effect of reading ability on talker-specific perceptual learning.

Note

This thesis reflects a working manuscript of a collaborative project conducted with Katlyn Salvador and Dr. Rachel M. Theodore. Ms. Salvador and I have worked together with respect to data collection and data analysis. This manuscript will be submitted for publication with authorship shared by all named above. My independent contributions to this project include writing the initial draft of the introduction and discussion sections as well as serving as primary coordinator for data collection.

Introduction

The acoustic signal of speech is a complex signal that simultaneously cues the linguistic content of a talker's message and the identity of the talker. In comprehension, the acoustic signal is the basis on which listeners access the sound structure of language in a process that leads to semantic comprehension. In addition to comprehension of a communicative message, listeners are able to identify specific talkers as well as myriad qualitative aspects about the talker including demographic differences, health status, age, gender, personality, and even emotional state of the talker (Murray & Arnott, 1993; Theodore, Miller, & DeSteno, 2009). Traditional models of speech perception proposed that the properties of the acoustic signal that cue linguistic meaning and talker identify were distinct (Halle, 1985). That is, it was believed that aspects of the signal that cued talker, personality, gender, or emotion were stripped away during the perception process in order to activate abstract linguistic units for comprehension (Licklider, 1952). However, a growing body of evidence challenges this traditional view (Mullenix, Pisoni, & Martin, 1989). Specifically, numerous findings indicate that linguistic and talker-specific phonetic information are tightly linked in language comprehension, at both early segmental levels of representation (Theodore & Miller, 2010) and at the lexical level of representation (Palmeri, Goldinger, & Pisoni, 1993).

Many findings indicate that listeners experience a host of comprehension benefits for familiar compared to unfamiliar talkers, suggesting that efficient language comprehension relies on integrating acoustic information about talker identity and the linguistic message (e.g., Goldinger, 1996). One of these benefits is increased intelligibility of the spoken message when listening to a familiar talker compared to an unfamiliar talker (Nygaard, Sommers, & Pisoni, 1994). Nygaard et al. (1994) trained participants to learn to recognize ten talkers over a nine-day training period. At the end of this nine-day period, listeners transcribed words produced by the talkers presented during training and novel talkers. The results showed that participants had increased word comprehension for the trained compared to novel talkers, indicating that voice recognition and processing of the phonetic content of a linguistic utterance are not independent. This talker familiarity benefit for intelligibility has been shown for both isolated words and sentences (Nygaard & Pisoni, 1998). The talker familiarity benefit generalizes to novel utterances, suggesting that listeners are learning about how talkers implement phonological segments and not specific words presented during training. Moreover, listeners show decreased processing time for familiar compared to unfamiliar talkers (Clarke & Garrett, 2004) and decreased resource allocation for familiar compared to unfamiliar talkers (Mullenix, Pisoni, & Martin, 1989).

This familiar talker benefit demonstrated in adults has also been shown in school-aged children (Levi, 2014). In this study, 41 native English-speaking children aged 7-12 were trained on 6 female bilingual German-English speakers for five days. A baseline measurement of spoken word recognition was obtained prior to training, which consisted of the child listening to a series of words (spoken by the 6 female voices they were to be trained on) and asked to repeat back the word that they heard. After explicit training on these 6 voices, children demonstrated a significant improvement in their ability to identify highly familiar words spoken by these familiar talkers (Levi, 2014). Additionally, children who performed the poorest at baseline (who were also the youngest subjects) demonstrated the greatest magnitude of improvement from pre- to post-test, indicating that talker familiarization may actually facilitate spoken language processing for children during this critical time of language acquisition.

Just as experience with a talker's voice influence linguistic processing, it has been shown that this link is bi-directional such that linguistic experience influence voice recognition abilities. Specifically, listeners show a native-language benefit for talker identification such that voice recognition is heightened for native compared to non-native talkers. Perrachione and Wong (2007) examined voice recognition for English and Mandarin speakers in two groups of listeners, native English listeners and native Mandarin listeners. Both groups of listeners competed a talker identification task and the results showed that English listeners were significantly better at identifying the English compared to the Mandarin talkers, but that the opposite pattern was found for the Mandarin listeners. The native-language benefit for voice recognition is also observed when the speech presented to listeners is played in reverse (Fleming et al., 2014), which suggests that the native-language benefit is mediated by sub-lexical information such as the sound structure of a language. This raises the possibility that impaired access to phonological knowledge might mediate voice recognition abilities.

Indeed, Perrachione, Del Tufo and Gabrieli (2011) demonstrated that adults with dyslexia show diminished voice recognition ability. Dyslexia is a specific learning disability that is neurobiological in origin, characterized by difficulties with accurate and/or fluent word recognition, and by poor spelling and decoding abilities (Lyon, 2003). It is a persistent difficulty, which is identified in childhood and continues through adulthood, despite adequate education and general intelligence (Ahissar, 2007). One hallmark deficit in this population concerns phonological tasks. Perrachione et al. (2011) found that adults with dyslexia demonstrated poor talker discrimination abilities for both native and non-native talkers, whereas non-impaired readers showed the native-language benefit for talker identification (Perrachione et al., 2011). This finding suggests that intact phonological abilities are responsible for the native-language talker identification benefit. Additional work in this domain suggests that the lack of a native-language benefit in dyslexic population may not be so absolute. Perea et al. (2014) extended these findings to compare differences between the developing system in children and the developed adult system. In separate experiments, adults and children with and without dyslexia participated in a talker identification task similar to that used in Perrachione et al. (2011). The results of this study indicated that both the adults and children with dyslexia showed a native-language benefit for talker identification (as did the control participants), but that overall performance for those with dyslexia was decreased relative to controls. These findings suggest that there may be a gradient influence of phonological ability on talker recognition, which is consistent with other findings showing such gradient influences including age of acquisition of a non-native language (Bregman & Creel, 2014). Indeed, recent findings from our laboratory (Kadam, Orena, Theodore, & Polka, in preparation) have demonstrated that reading ability exerts a gradient influence on native and non-native talker identification among unimpaired readers.

Viewed collectively, these studies indicate that experience with a talkers' voice facilitates language comprehension, and that phonological knowledge influence voice recognition. What is not known, however, is whether the phonological factors that influence voice recognition ability will also influence the degree to which listeners achieve talker familiarity effects for language comprehension. That is, will reading ability show a gradient effect on the ability to not only learn talkers' voices, but also the ability to incorporate talker-specific phonetic detail into lexical representations? One striking finding in the literature on talker-contingent speech perception is that not all learners are equally good learners. Specifically, Nygaard & Pisoni (1998) separated participants into two groups based on performance during the talker identification task. Good learners were designated as those who were able to identify talkers with 70% accuracy on the last day of training, with the rest of the participants designated as poor learners. Learning groups differed in three ways. First, “good learners” demonstrated higher talker identification accuracy overall, not just for the last day of training on which the learning split was determined. Second, “good learners” had a larger magnitude of learning despite having comparable performance to “poor learners” on the first day of training, indicating that the good learners had a higher rate of learning compared to the poor learners. Finally, “good learners” showed higher performance on the speech perception task at post-test compared to the “poor learners.” That is, those who excelled at the talker identification task received the greatest benefit of talker familiarity with respect to language comprehension.

What this study did not reveal, however, was the specific linguistic, cognitive, or auditory abilities that differed between the good and poor learners and thus was responsible for the differences between the groups. Here we examine one possibility, which is that reading ability may influence the talker familiarity effect for language comprehension. Efficient reading is the output of a host of processes, many of which are phonological in nature (Peterson & Pennington, 2012). Indeed, the central deficit in reading disability is an impairment in phonological awareness and rapid naming abilities (Lyon, Shaywitz, & Shaywitz, 2003), both of which are important for fluent decoding of print leading to better comprehension. Research has not only shown that those with dyslexia show impaired talker recognition abilities (e.g., Perrachione et al., 2011), but also that reading ability exerts a gradient influence on talker identification among unimpaired readers (Kadam et al., in press). If the ability to learn a talker’s voice influences the degree to which talker-specific lexical processing will occur, as suggested by Nygaard & Pisoni (1998), then we predict that highly skilled readers will show heighted talker-specific perceptual learning compared to average readers. A failure to observe such a difference would suggest that the locus of the talker-specific learning benefit is not phonological in nature.

Methods

Participants

Thirty adults were recruited from the University of Connecticut community for participation. The participants (11 males, 19 females) were between 18 and 23 years of age (M = 20, SD = 2). All participants were native, monolingual speakers of American English with no history of speech, language, or hearing disorders according to self-report. Three participants reported a previous history of reading difficulty, but (as described in detail below) did not meet criterion for reading disability based on current performance on a standardized assessment battery. All participants passed a pure tone hearing screen on the first day of participation administered at 20 dB for octave frequencies between 500 and 4000 Hz. Participants were either paid or received partial course credit for their participation. All testing procedures and acquisition of informed consent followed protocols approved by the University of Connecticut Institutional Review Board.

All participants completed a standardized assessment battery in order to examine nonverbal intelligence, memory, reading sub-skills, and reading comprehension. The components of the battery were the Test of Nonverbal Intelligence, 4th Edition (TONI – 4); the Auditory Memory Index (AMI) and Immediate Memory Index (IMI) of the Wechsler Memory Scale, 4th Edition (WMS – IV); the Elision, Blending Words and Nonword Repetition subtests of the Comprehensive Test of Phonological Processing, 2nd Edition (CTOPP – 2); the RAN Numbers, RAN Letters, and RAS 2 – Set Letters and Numbers subtests of the Rapid Automatized Naming and Rapid Alternating Stimulus Tests (RAN/RAS); the Sight Word Efficiency and Phonemic Decoding Efficiency subtests of the Test of Word Reading Efficiency, 2nd Edition (TOWRE – 2); and the Word Identification, Word Attack, and Passage Comprehension subtests of the Woodcock Reading Mastery Test, 3rd Edition (WRMT – III).

Participants were assigned to either the average reading group or the advanced reading group based on performance for the reading sub-skills and reading comprehension measures. Specifically, we first calculated a composite reading score for each participant defined as the mean percentile across the standardized measures relevant to reading sub-skills and reading comprehension. Then, we performed a median split across the participants, with the lower half (n = 15) assigned to the average reading group and the upper half (n = 15) assigned to the advanced reading group. The mean composite reading score for the average readers was 57 (*SD* = 9), which was statistically lower than the mean composite reading score for the advanced readers 77 (*SD* = 3), t (28) = -8.051, p = 0.000, d = -2.940. As shown in Table 1, which reports mean performance and test statistics for the two reading groups on each of the standardized measures, the difference in performance between the two groups for the composite scores is consistent with the difference for the individual measures. Critically, the two groups did not differ with respect to performance on cognitive and memory assessments (i.e., TONI – 4, WMS – III).

Stimuli

The stimulus set consisted of auditory recordings of 232 words produced by six talkers (3 male, 3 female) and six cartoon faces, each one paired with a specific talker. Lexical characteristics for each item were determined using the MRC Psycholinguistic Database (Wilson, 1988) and CLEARPOND (Marian et al., 2012). The words consisted of 116 one-syllable items and 116 two-syllable items. For each syllable type, half of the words were selected to be high frequency items (*M* = 100, *SD* = 260) and the other half to be low frequency items (*M* = 6, *SD* = 4) based on the verbal frequency counts reported in the MRC Psycholinguistic

Database. For the one-syllable words, neighborhood density (low versus high) was manipulated with respect to word frequency, such that half of the one-syllable words could be considered "easy" words (i.e., high frequency items with low neighborhood density) and the other half could be considered "hard" words (i.e., low frequency with high neighborhood density). The criterion for marking low versus high neighborhood density was 20 lexical neighbors (i.e., fewer than 20 neighbors was categorized as low density). The 232 words were randomly assigned to one of three lists, with the constraint that each list was balanced with respect to the lexical characteristics described above. Specifically, four items were assigned to the familiarization list, 114 items were assigned to the training list (which was used during talker identification training and the pre- and post-tests), and 114 items were assigned to the novel list (for use only during the pre- and post-tests).

Acoustic recordings of the 232 words were obtained from six talkers (3 males and 3 females). The six talkers were fictitiously referred to as Matt, Paul, Zack, Anne, Beth, and Jill. Each talker was recorded producing multiple repetitions of each word along with filler items. All recordings took place in a sound-attenuated booth. Speech was recorded via microphone (AKG D5) connected to pre-amplifier (Digidesign MBox 2) and saved directly to hard disk using the Praat software (Boersma & Weenink, 2011). Each word was excised from the longer recording and saved to an individual file using Praat. One repetition of each word that was intelligible and free of acoustic artifact was selected for use in the stimulus set. Finally, the 1392 selected tokens (232 words X 6 talkers) were equated for root-mean-square amplitude. In order to describe the variation among the selected talkers, we calculated mean fundamental frequency, jitter, shimmer, and word duration (as a metric of speaking rate) for each talker; these values are shown in Table 2. Note that the three voice measurements were determined for a 100 ms portion of each token that was centered at the peak amplitude of that token. Each talker was assigned to a cartoon face, as shown in Figure 2.

Procedure

The overall procedure consisted of a word identification pre-test, three sessions of talker identification training, and a word identification post-test. The three training sessions were always administered on three different days, and the pre- and post-tests were always separated by the talker identification training. The mean time to complete all sessions across the 30 participants was 5 days (*SD* = 1). Specific details for the test and training sessions are described below.

Pre-test and post-test. The procedure for the pre- and post-tests was identical. All testing took place in a sound-attenuated booth. Participants were seated at a table that contained a response box, a computer keyboard, and a computer monitor. Auditory stimuli were presented via headphones (Sony MDR-7506) in the context of white noise low pass filtered at 4800 Hz. Both speech and noise were presented at 65 dB SPL, yielding a S/N of 0 dB. Responses were collected via the computer keyboard.

The stimuli consisted of one production of each of the 228 lexical items from the training and novel lists, with 38 items produced by each of the six talkers. The lexical characteristics of the items for a given talker were balanced across the set of talkers for both the training and novel items. One randomization of the 228 items was presented. For each trial, the participant was directed to transcribe the word using the computer keyboard and then press the enter key to begin the next trial. Participants were allowed to edit their response prior to pressing the enter key. If a response was not collected within 7 seconds, then the next trial began. No feedback was provided at test. Each test lasted approximately 25 minutes.

Talker-training. Three talker-training sessions were administered on different days intermediate to the pre-test and post-test. All training sessions took place in a sound-attenuated booth. Participants were seated at a table that contained a response box and a computer monitor. Auditory stimuli were presented via headphones (Sony MDR-7506) at 65 dB SPL without the presence of background noise. Responses were collected via the response box. The stimuli consisted of a separate training list for each of the three training sessions. Across the three training sessions, participants heard each talker produce each word one time. Each training list consisted of 228 items, which was two repetitions of the 114 training words, each repetition produced by a different talker. Accordingly, each talker contributed 38 items to each of the 3 training lists.

Participants were instructed to watch and listen to a brief familiarization phase prior to each training session in order to begin to associate a specific talker's voice (presented via headphones) with its corresponding cartoon face and name that were simultaneously displayed on the monitor. The four familiarization items (described above) were presented for each of the six talkers, one talker at a time. The same four words were used for every talker, and both the words and the talkers were always presented in the same order. This phase took approximately 2 minutes and the procedure was identical for each of the three sessions.

After the familiarization phase was completed, the participant began the talker-identification training. For each trial., a word was presented via headphones simultaneously with the face array shown in Figure 1. The participant was directed to indicate which cartoon face matched the voice of the auditory word by pressing an appropriately labeled button on a response box. Feedback was provided after every trial in the form of "YES" or "NO" appearing on the computer monitor for correct and incorrect responses, respectively. Each training session lasted approximately 25 minutes.

Results

Training

For each participant, mean percent correct talker identification was calculated separately for each training session, collapsing over the six talkers presented during training. Mean percent correct talker identification for the average and advanced reading groups is shown in Figure 1 for each training session. Visual inspection of this figure suggests that performance improved over the three training sessions for each reading group, but that the rate of talker learning was faster for the advanced readers compared to the average readers. Specifically, it appears that talker identification was equivalent for the two reading groups during the first and last training sessions, but that performance differed between the two groups during the second training session, with talker identification improved for the advanced compared to the average readers.

To examine these patterns statistically, we submitted mean percent correct talker identification to a mixed-ANOVA with the within-subjects factor of training session (session 1, session 2, session 3) and the between-subjects factor of reading group (average and advanced). The results showed a main effect of training session, $F\ (2,56) = 67.92$, $p < 0.001$, $\eta_p^2 = 0.708$, a marginal effect of reading group, $F\ (1,28) = 3.57$, $p < 0.069$, $\eta_p^2 = 0.708$, and a reliable interaction between training session and reading group, $F\ (2,56) = 4.01$, $p = 0.024$, $\eta_p^2 = 0.125$. To explicate the nature of the interaction, we used independent t-tests to compare performance between the two reading groups for each training session. There were no differences between the two reading groups at session 1, $t\ (28) = -0.72$, $p = 0.478$, $d = -0.272$, or session 3, $t\ (28) = -1.42$, $p = 0.166$, $d = -0.536$, but the average readers ($M = 74$, $SD = 7$) showed lower talker identification accuracy compared to the advanced readers ($M = 86$, $SD = 13$) at session 2, $t\ (28) = -3.11$, $p = 0.004$, $d = -1.175$.

Thus, it appears that the interaction between training session and reading group emerged not due to a difference in performance at the end of the training period, but rather a difference with respect to how much training was need to meet that level of talker identification. This interpretation was confirmed with a set of paired t-tests that examined performance across the three sessions within each reading group. The average readers showed significant improvement in talker identification from session 1 to session 2, t (14) = -4.60, $p < 0.001$, d = -1.188, and from session 2 to session 3, t (14) = -3.11, p = 0.008, d = -0.802. The advanced readers, however, showed striking improvement from session 1 to session 2, t (14) = -5.64, $p < 0.001$, d = -1.457, with no additional gains observed from session 2 to session 3, t (14) = -1.15, p = 0.271, d = -0.296.

Test

Performance at test was measured in terms of percent correct word transcription. A response was considered correct if the participant's keyed entry matched the correct spelling. A response was also considered correct if the participant's keyed entry was an obvious typing error or a common misspelling, with the constraint that the participant's entry was not counted correct if the alternate spelling created a different word. For each participant, mean percent correct word transcription was calculated separately for training and novel words during the pre-test and post-test.

Mean percent correct intelligibility for the 30 participants was submitted to ANOVA with the within-subjects factors of time (pre- vs. post-test) and trial type (trained vs. novel items) and the between subjects factor of reading group (average vs. advanced). The results of the ANOVA revealed a main effect of time, F (1,28) = 39.83, $p < 0.001$, $\eta_p^2 = 0.587$, with performance at post-test higher than performance at pre-test. The ANOVA also revealed a marginal main effect of reading group, F (1,28) = 4.13, p = 0.052, η_p^2 = 0.128, with performance of the advanced readers higher compared to performance of the average readers. Critically, the ANOVA showed a significant interaction between test time and reading group, F (1,28) = 7.784, p = 0.009, η_p^2 = 0.218.

Figure 3 shows mean percent correct word transcription for the two reading groups at pre- and post-test. Visual inspection of the figure suggests that the advanced readers showed increased learning at post-test compared to the average readers. This pattern was confirmed with independent t-tests that showed no difference in word transcription between the two groups at pre-test, t (28) = -0.41, p = 0.684, d = -0.154, and significantly lower word transcription for the average compared to the advanced readers as post-test, t (28) = -3.44, p = 0.002, d = -1.300. Paired t-tests confirmed that both the average, t (14) = -2.50, p = 0.026, d = -0.645, and advanced readers, t (14) = -6.310, p < 0.001, d = -1.631, improved from pre- to post-test. Thus the interaction between reading group and test time was not a consequence of the average readers not improving; rather, it was a consequence of the advanced readers improving more compared to the average readers.

The ANOVA also showed a main effect of trial type, F (1,28) = 61.31, p < 0.001, η_p^2 = 0.686, with performance improved for the trained items compared to the novel items, and an interaction between trial type and test time, F (1,28) = 35.793, p < 0.001, η_p^2 = 0.561. However, there was no interaction between trial type and reading group, F (1,28) = 0.02, p = 0.878, η_p^2 = 0.001, indicating that generalization was equivalent for the average compared to the advanced readers. The three-way interaction between trial type, test time, and reading group, F (1,28) = 1.82, p < 0.188, η_p^2 = 0.061, was not significant.

One final analysis was conducted in order to determine if the patterns observed between the two reading groups would be observed if reading ability were considered in a continuous fashion, rather than as a categorical grouping. For this analysis, we considered performance at pre- and post-test for all 30 participants as a function of their composite reading score (i.e., not as the categorical grouping of average vs. advanced readers). Figure 4 shows the correlation between both pre- and post-test word transcription across the range of reading ability measured in our sample. The correlation between transcription accuracy and composite reading ability was not significant at pre-test, $r = 0.28$, $p = 0.134$, but was at post-test, $r = 0.56$, $p = 0.001$. This result suggests that there is a gradient influence of reading ability on talker-specific perceptual learning such that as abilities that influence reading ability increase, so too does the ability to modify speech perception as a consequence of experience with talkers' voices.

Summary and Conclusions

The acoustic signal of speech provides a rich source of information cueing semantic meaning of the spoken message and the identity of the talker. Research has demonstrated that there is a comprehension benefit when listening to a familiar talker compared to an unfamiliar talker (Clarke & Garrett, 2004; Levi, 2014; Nygaard et. al., 1994; Nygaard et al., 1998) Research has also demonstrated that linguistic competence facilitates talker recognition (Perrachione & Wong, 2007; Bregman & Creel, 2014). These findings from a healthy perceptual system have since been extended to an impaired population, specifically, individuals with dyslexia. Results showed significant group differences between individuals with and without dyslexia in both adults (Perrachione et al., 2011) and children (Perea et al., 2014). Previous work indicates group differences between an intact and impaired system, specifically with respect to reading ability. Previous work also indicates that unimpaired individuals show graded learning ability for talker identification that has yet to be fully explicated (Nygaard et al., 1998). In developing a more concrete model of spoken language processing that accounts for the links between talker identification abilities and speech perception abilities, the current work examined the role of reading ability on talker-specific perceptual learning. We predicted that average readers, those who performed around the center of the normal distribution on a diagnostic battery, would demonstrate different abilities in talker identification and speech perception tasks than advanced readers who scored in the top portion of the normal distribution on the same battery. Specifically, we predicted that individuals with advanced reading abilities would demonstrate a greater ability to both learn talkers' voices and incorporate talker-specific sub-lexical information into their perceptual systems.

Results yielded significant findings for both training and test portions of this experiment. In training, we found a robust difference in talker learning rate between the average and advanced readers. Though both groups of readers began and ended the talker training sessions with comparable accuracy, the advanced readers reached that level in two training sessions whereas the average readers required three training session to reach equivalent learning. At test, we observed an interaction between reading ability and performance on the pre- and post-test. Though both groups showed evidence of talker familiarity effects on word recognition, the learning effect was greater in the advanced compared to the average readers. This relationship between reading ability and talker-specific perceptual learning held not just for the categorical grouping of average versus advanced readers, but also was observed across the continuous measure of reading ability.

The current findings contribute to the broader idea that there is something inherently different about an individual's linguistic architecture facilitating language expertise that allows for better performance than others on talker identification. It is possible that the same aspect of the language system that promotes efficient word recognition also promotes efficient voice recognition. If this is about gradient effects of language experience on talker identification, perhaps phonological knowledge is what is providing this advantage. This would support previous findings indicating that phonological knowledge contributes to graded talker identification abilities (Perrachione & Wong, 2007; Bregman & Creel, 2014). Given that the present study is measuring how well people modify phonological knowledge, something about having stronger performance on this reading diagnostic battery allows an individual to better take advantage of acoustic input during training. Being trained to identify a talker's voice provides a perceptual advantage in comprehension, but as results from this study indicate, advanced readers are able to take better advantage. These findings compliment prior findings indicating group differences between individuals with and without dyslexia (Perrachione et al., 2011; Perea et al., 2014) by extending findings across the normative distribution of reading abilities.

The current work provides critical data in moving towards a perceptual model that accounts for how language ability and talker identification are both influencing each other and constraining one another. In considering future directions of the current work, further research is needed to explore this relationship. Perhaps it is the case that within the linguistic framework, there is an unidentified link that is equally contributing to talker identification abilities and comprehension abilities in individuals across the range of reading abilities, rather than these two processes directly influencing one another. Such possibilities could be confirmed by examining structural and functional differences between the neuroanatomy of individuals with dyslexia and typically developing individuals. Researchers consistently observed decreased white matter in the areas of the left temporoparietal regions and the left inferior frontal gyrus, two areas of the brain that are highly correlated with reading skill (Peterson & Pennington, 2012). Other studies have found that individuals with dyslexia have greater than typical white matter in the corpus collosum, suggesting that these pathways that support reading project too weakly in the linguistic left hemisphere, and project too strongly into the right hemisphere, which is less efficient for linguistic processing (Gabrieli, 2009). Functional differences have also been demonstrated. During reading tasks, individuals with dyslexia demonstrate consistent under activation in areas of the brain known to aid in phonological processing and reading tasks, including the left temper parietal region, the left temporoparietal region, the left inferior frontal gyrus, and the left occipitotemporal region (Gabrieli, 2009; Pugh et al., 2000; Shaywitz et al., 2001, Shaywitz & Shaywitz, 2005). It is possible that these structural and functional differences that exhibit significant differences among the impaired and unimpaired population will show graded differences when considering the entire normative curve of readers. For example, perhaps it is the case that those who are identified as average readers have greater than typical white matter in the corpus collosum than individuals who are advanced readers, but less projections than those with reading impairment or dyslexia.

Further research is also needed when drawing conclusions about the relationship between talker identification and speech perception in considering research in the adult compared to the developing child system. It is possible that a lifetime of experience with advanced, average, or impaired reading ability exerts a graded influence on one's ability to perform on tasks, rather than these graded abilities existing at onset. Further research in children, and specifically, children before they learn to read, could contribute to this growing body of evidence.

References

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11*(11), 458-465.

Ahissar, M. (2006) Dyslexia and the failure to form a perceptual anchor.(2006). *Nature Neuroscience, 9*(12), 1558-1564.

Boersma, Paul & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.45, retrieved 2011 from http://www.praat.org/

Bregman, M. R., & Creel, S. C. (2014). Gradient language dominance affects talker learning.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116*(6), 3647-3658.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*(2), 224-238.

Gabrieli. J.D. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science, 325*(5938), 280-283.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166-1183.

Halle, M. (1985). Speculations about the representation of words in memory. (pp. 101-114)

Kadam, M., Orena, A. J., Theodore, R. M., & Polka, L. (In press). Gradient effects of reading ability on native and non-native talker identification. *Proceedings of the 18th International Congress of Phonetic Sciences.*

Levi, S. V. (2014). Talker familiarity and spoken word recognition in school-age children. Journal of Child Language, 1-30.

Licklider, J. C. R. (1952). On the process of speech perception. *The Journal of the Acoustical Society of America, 24*(6), 590.

Lyon G., Shaywitz, S.A., & Shaywitz, B.A. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading: A definition of dyslexia. *Annals of Dyslexia, 53*, 1.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85*(1), 365-378.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93*(2), 1097-1108.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60*(3), 355-376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*(1), 42-46.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 309-328.

Perea, M., Jimenez, M., Suarez-Coalla, P., Fernandez, N., Vina, C., & Cuetos, F. (2014). Ability for voice recognition is a marker for dyslexia in children. *Experimental Psychology, 61*(6), 480-487.

Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. E. (2011). Human voice recognition depends on language ability. *Science, 333*(6042), 595-595.

Perrachione, T. K., & Wong, P. C. M. (2007). Learning to recognize speakers of a non-native language: Implications for the functional organization of human auditory cortex. *Neuropsychologia, 45*(8), 1899-1910.

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *Lancet, 379*(9830), 1997.

Pugh, K. R., Fletcher, J. M., Liberman, A. M., Shankweiler, D. P., Katz, L., Lacadie, C., et al. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *Psychological Science, 11*(1), 51-56.

Shaywitz, B. A., & Shaywitz, S. E. (2005). Dyslexia (specific reading disability). *Biological Psychiatry, 57*(11), 1301-1309.

Shaywitz, B. A., Shaywitz, S. E., Gore, J. C., Pugh, K. R., Fulbright, R. K., Mencl, W. E., et al. (2001). The neurobiology of dyslexia. *Clinical Neuroscience Research, 1*(4), 291-299.

Theodore, R. M., & Miller, J. L. (2010). Characteristics of listener sensitivity to talker-specific phonetic detail. *Journal of the Acoustical Society of America, 128*(4), 2090.

Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *Journal of the Acoustical Society of America, 125*(6), 3974-3982.

Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Pedrotty Bryant, D., Dickson, S., & Blozis, S. A. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education, 24*(5), 301-315.

Vellutino, F. R., Pruzek, R. M., Steger, J. A., & Meshoulam, U. (1973). Immediate visual recall in poor and normal readers as a function of orthographic-linguistic familiarity. *Cortex, 9*(4), 370-386.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*(3), 415-438.

Table 1

Mean, standard deviation (in parentheses), *t*, *p*, and Cohen's *d* for the average and advanced readers for each component of the standardized assessment battery and the composite reading score. The *t* and *p* values reflect those derived from independent t-tests (*df* = 28) for each assessment measure. See the main text for a description of each assessment.

Assessment		**Average Readers**	**Advanced Readers**	***t***	***p***	***d***
TONI–4		41 (27)	33 (23)	0.838	0.409	0.306
WMS–4	AMI	53 (21)	54 (21)	-0.097	0.924	-0.036
	IMI	57 (21)	60 (25)	-0.401	0.692	-0.146
CTOPP	Elision	48 (19)	64 (13)	-2.737	0.011	-0.999
	Blending	52 (29)	79 (12)	-3.395	0.002	-1.240
	Nonword Repetition	49 (25)	55 (24)	-0.763	0.452	-0.279
RAN/RAS	RAN Numbers	71 (16)	80 (6)	-2.193	0.037	-0.801
	RAN Letters	63 (16)	75 (7)	-2.748	0.010	-1.003
	RAS 2-Set	71 (8)	84 (6)	-5.019	0.000	-1.833
TOWRE	Sight Word	64 (25)	83 (16)	-2.500	0.019	-0.913
	Phonemic Decoding	57 (20)	88 (8)	-5.683	0.000	-2.075
WRMT–III	Word Identification	43 (19)	83 (15)	-6.399	0.000	-2.337
	Word Attack	41 (25)	77 (14)	-4.828	0.000	-1.763
	Passage Comprehension	67 (17)	78 (12)	-2.029	0.052	-0.741
Composite Reading Score		57 (9)	77 (3)	-8.051	0.000	-2.940

Table 2

In hertz, mean average fundamental frequency (mean F_0), mean minimum fundamental frequency (min F_0), mean maximum fundamental frequency (max F_0), and mean standard deviation of fundamental frequency (SD F_0) for the six talkers used in the current study. Standard deviation for each metric is shown in parentheses.

Talker	Mean F_0	Min F_0	Max F_0	SD F_0
Anne	196 (20)	141 (50)	246 (25)	31 (16)
Beth	220 (8)	196 (30)	254 (16)	13 (8)
Jill	182 (8)	153 (32)	213 (19)	16 (10)
Matt	111 (8)	97 (7)	145 (37)	12 (12)
Paul	110 (7)	101 (9)	124 (32)	6 (11)
Zack	91 (8)	79 (3)	112 (27)	10 (9)

Figure 1

The six cartoon faces and talker names used in the talker identification training task. During training, the six faces appeared in a fixed horizontal array, with the three faces in the bottom row appearing to the right of the three faces in the top row.

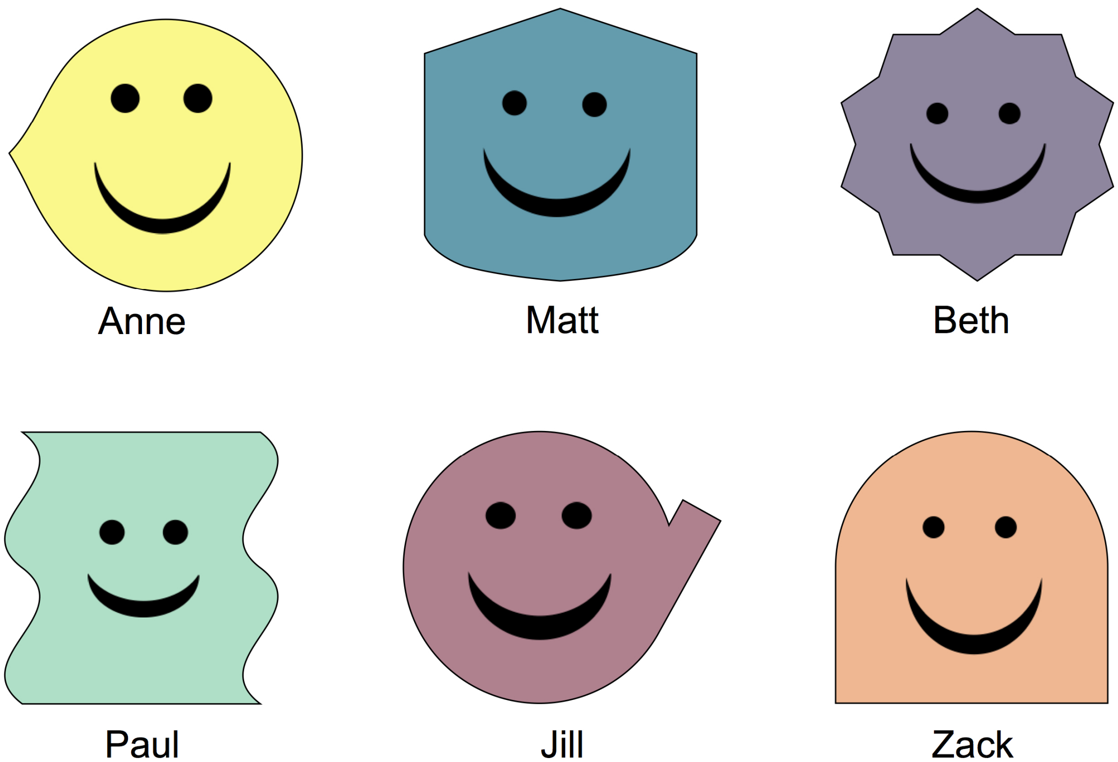

Figure 2

Mean talker identification accuracy for the average and advanced reading groups for each day of talker training. Error bars indicate standard error of the mean.

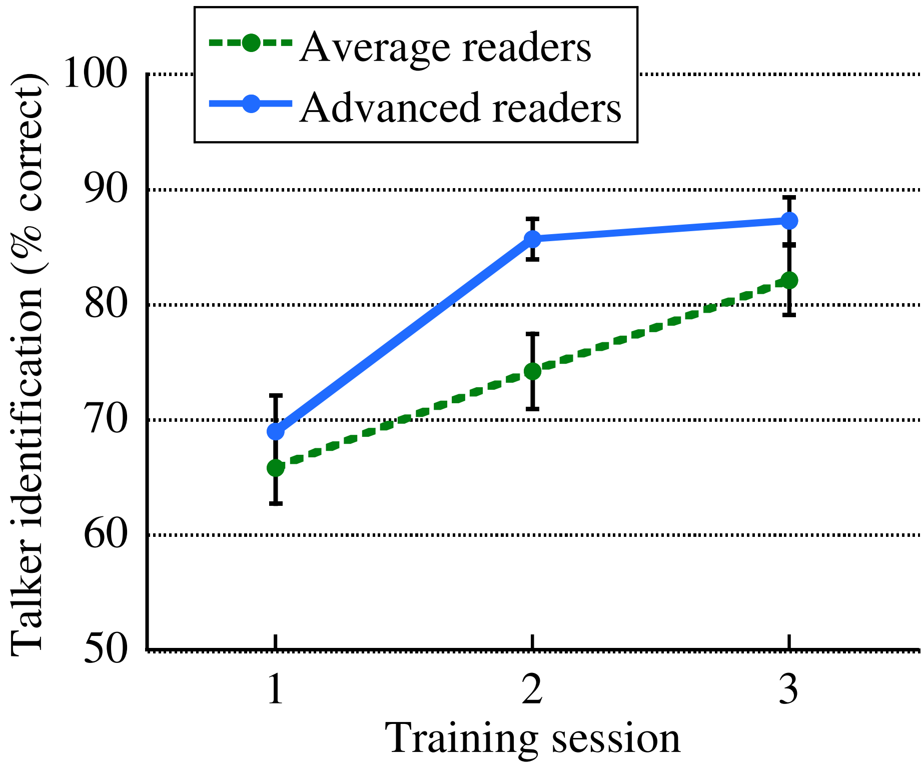

Figure 3

Mean word identification accuracy for the average and advanced reading groups at pre-test and post-test. Error bars indicate standard error of the mean.

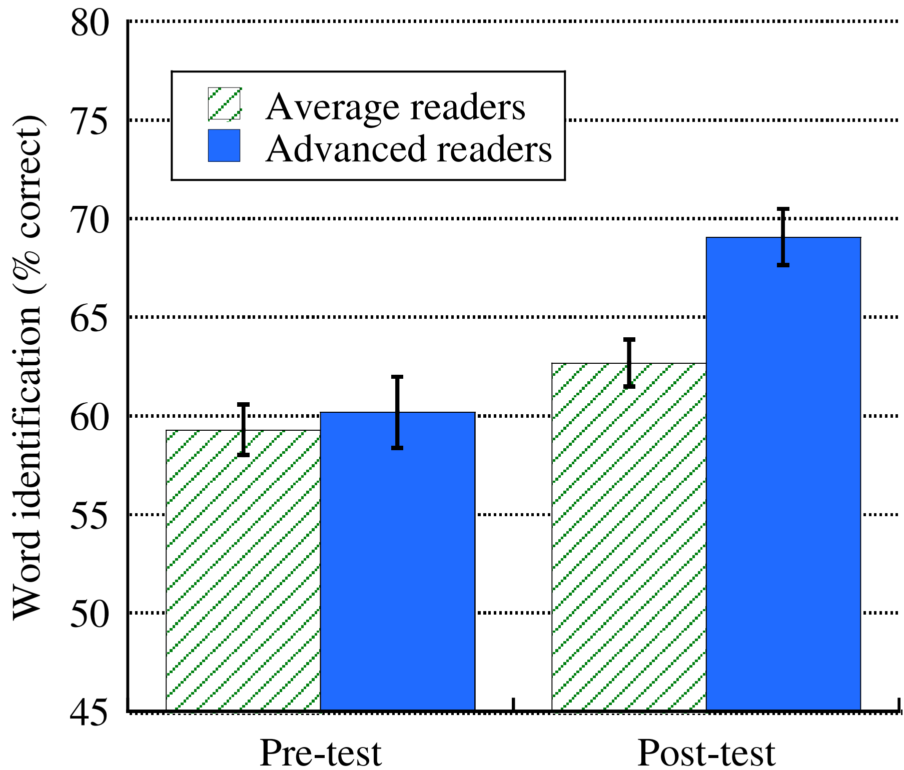

Figure 4

Relationship between word identification and composite reading score at pre-test and post-test. Each line shows the linear fit between word identification and composite reading score and Pearson's *r* for each line shown in the legend.